AD_____

Award Number: DAMD17-97-1-7283

TITLE: The Roles of Chromosome Breaks and Telomere Dynamics in the Genomic Instability Associated with Human Breast Cancer

PRINCIPAL INVESTIGATOR: John H. Wilson, Ph.D.

CONTRACTING ORGANIZATION: Baylor College of Medicine
Houston, Texas 77030-3498

REPORT DATE: September 2000

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	September 2000	Final (11 Aug 97 - 10 Aug 00)

4. TITLE AND SUBTITLE
The Roles of Chromosome Breaks and Telomere Dynamics in the Genomic Instability Associated with Human Breast Cancer

5. FUNDING NUMBERS
DAMD17-97-1-7283

6. AUTHOR(S)
John H. Wilson, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Baylor College of Medicine
Houston, Texas 77030-3498

E-MAIL:
jwilson@bcm.tmc.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 Words)*

As part of their progression from normal to malignant cells, human tumors acquire a marked genomic instability, which is likely due in part to the progressive shortening and transient loss of telomeres from chromosome ends. Loss of telomeres allows chromosomes to fuse end-to-end, triggering chromosome fusion-bridge-breakage cycles that lead to genome rearrangements, loss of heterozygosity, and gene amplification. Our objective was to study the initial steps in this process using site-specific double strand breaks (DSBs) adjacent to interstitial telomere sequence (ITS) and a color-based detection system on a specially engineered chromosome. Over the course of this grant we have determined the inherent instability of ITS, shown that a DSB on the chromosome can lead to the seeding of new telomeres, defined the orientation of the APRT gene on the chromosome, identified chromosome loss as a problem, redesigned the test chromosome to avoid chromosome loss, constructed a new targeting vector, prepared an appropriate recipient cell line, developed the color-based detection system, and created a cell line carrying the redesigned chromosome. Characterization of the new cell line will set the stage for many of the experiments proposed in the original application; we will pursue those studies with funding currently being requested from other sources.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Breast Cancer	58
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

Cover..... 1

SF 298..... 2

Table of Contents..... 3

Introduction..... 4

Body..... 4

Key Research Accomplishments.....12

Reportable Outcomes.....12

Conclusions.....12

References.....13

Appendices.....14

Bibliography.....15

List of Personnel.....15

Attachments (Appendices 1 & 2)

INTRODUCTION

The progression of tumors from benign to malignant is often marked by genomic instability, resulting commonly in a highly rearranged genome. It has recently been shown that progressive shortening and transient loss of telomeres from the ends of chromosomes may contribute significantly to these genome rearrangements (1). Loss of telomeres allows chromosomes to fuse end-to-end (2,3), initiating cycles of chromosome fusion-bridge-breakage (4,5), which can result in genome rearrangements, gene amplification, and loss of heterozygosity (6,7). The purpose of the research supported by this grant was to define the lengths of telomere sequence sufficient to prevent such rearrangements and to study the initial steps in chromosome fusion-bridge-breakage cycles. These processes were proposed to be investigated by introducing site-specific double-strand breaks (DSBs) adjacent to interstitial telomere sequences (ITSs) in a genetically marked region of a specially engineered test chromosome. Ultimately, the test chromosome, which was to be developed in Chinese hamster ovary (CHO) cells, would be transferred into breast cancer cells at different stages of tumor progression to assess frequencies of fusion-bridge-breakage cycles in those cells.

BODY

The scope of the proposed research is summarized by the technical objectives, as outlined below:

Technical Objective 1. *Measure frequencies of chromosome loss after introduction of a site-specific double-strand break by I-SceI expression.*

Technical Objective 2. *Use microcell fusion to transfer the test chromosome to a breast cancer cell line and repeat measurements of effects of different telomere lengths and telomerase expression.*

Technical Objective 3. *Construct a color-based system for detecting chromatid fusion and gene amplification.*

Technical Objective 4. *Characterize the color-based system in CHO cells for its ability to detect chromatid fusion and gene amplification.*

Technical Objective 5. *Transfer the color-based detection system into breast cancer cells at different stages in the progression to malignancy and measure frequencies of chromatid fusion and gene amplification.*

As detailed below, we have made substantial progress in our understanding of ITS stability, of the requirements for a test chromosome, and in the development of a color-based detection system. We have also encountered unanticipated problems that have necessitated a redesign of essential components of our test system. That has been completed and a cell line has been constructed and is being characterized. The resultant time delay has prevented our accomplishing all the stated objectives of the original application. We have, however, laid a firm experimental foundation for the stated objectives, which are still highly relevant, and we anticipate obtaining future funding to pursue them. Below we discuss the progress that has been made on this grant.

Instability of interstitial telomere sequence.

In work that was begun before this grant was initiated, but completed over the course of the grant, we investigated the inherent instability of ITS. To support our proposed experiments, an ITS must be relatively stable. Telomere-repeat sequences cap the ends of eucaryotic chromosomes and help stabilize

them (8). At interstitial sites, however, they may destabilize chromosomes, as suggested by cytogenetic studies in mammalian cells that correlate interstitial telomere sequence with sites of spontaneous and radiation-induced chromosome rearrangements (9-17). In no instance is the length, purity, or orientation of the telomere repeats at these potentially destabilizing interstitial sites known. To determine the effects of a defined interstitial telomere sequence on chromosome instability, as well as other aspects of DNA metabolism, we deposited 800 bp of the functional vertebrate telomere repeat, TTAGGG, in two orientations in the second intron of the adenosine phosphoribosyltransferase (APRT) gene in Chinese hamster ovary cells. In one orientation the deposited telomere sequence did not interfere with expression of the APRT gene, whereas in the other it reduced mRNA levels by about 50%. The telomere sequence did not induce chromosome truncation and the seeding of a new telomere at a frequency above the limits of detection. Similarly, the telomere sequence did not alter the rate or distribution of homologous recombination events. The interstitial telomere repeat sequence in both orientations, however, dramatically increased gene rearrangements some 30-fold. Analysis of individual rearrangements confirmed the involvement of the telomere sequence. These studies define the telomere repeat sequence as a destabilizing element in the interior of chromosomes in mammalian cells. The level of this instability, although dramatically elevated over background, is still low enough that it poses no difficulties for the proposed experiments.

This work is described in detail in Appendix 1, which is a manuscript that is under review for *Molecular and Cellular Biology.*

Seeding a new telomere by the joining of a telomere sequence to a double-strand break.

A critical component of our general approach to the proposed studies was that interstitial telomere sequence could lead to the seeding of new telomeres. It was crucial that we show that the telomere sequence we proposed to use could, in fact, support the seeding of new telomeres in CHO cells. In work that was initiated before this grant began, we constructed a targeted CHO cell line that carries the recognition site for the rare cutting endonuclease, I-SceI, in the second intron of the APRT gene and characterized the effects of cutting. By co-transfecting an I-SceI expression vector and a selectable plasmid bearing telomere sequence, we sought to select for integration and telomere seeding at a defined chromosomal break. Before we could apply this strategy, however, it was necessary to create tetraploid versions of this cell line so that any essential DNA in the largely haploid region surrounding the APRT locus would be rendered nonessential. Tetraploid cells were created by fusing the I-SceI-containing, APRT+ cell line with a cell line that is deleted for the APRT gene (18,19).

Diploid and tetraploid cell lines containing the I-SceI site at the APRT locus were transfected individually or cotransfected with the I-SceI expression vector and the linearized plasmid pSXNeoTTAGGG, which contains telomere sequence fused to a neomycin expression cassette. Transfected cells were selected for Neo+, for APRT-, and for Neo+APRT- cells. The data are presented in Table 1.

Table 1. Transfection of diploid and tetraploid cell lines with a plasmid expressing I-SceI, or a telomere-containing plasmid expressing Neo+, or both plasmids.

Transfected DNA	**Neo+**		**APRT-**		**Neo+APRT-**	
	Diploid	**Tetraploid**	**Diploid**	**Tetraploid**	**Diploid**	**Tetraploid**
pCMVI-SceI	NA	NA	$3x10^{-4}$	$7x10^{-4}$	NA	NA
pSXNeo	$6x10^{-3}$	$2x10^{-3}$	$5x10^{-6}$	$8x10^{-5}$	$<2x10^{-6}$	$9x10^{-5}$
pCMVI-SceI + pSXNeo	$5x10^{-3}$	$1x10^{-3}$	$5x10^{-4}$	$4x10^{-4}$	$3x10^{-4}$	**$1x10^{-2}$**

The data for Neo+ alone indicated that I-SceI cleavage does not significantly alter the baseline frequency of random integration in either diploid or tetraploid cells; however, it did indicate that the tetraploid cell line gave 3- to 5-fold fewer random integrants than the diploid cell line. The APRT- data showed that in the diploid cells, I-SceI cleavage stimulates formation of APRT- colonies about 100-fold, in agreement with our published data. Although the stimulated values were about the same in diploid and tetraploid cells, in tetraploid cells it represented only about a 5- to 10-fold stimulation over the frequency in the absence of I-SceI expression. The 10-fold higher background of APRT- cells in tetraploid cells likely represents the frequency of loss of the APRT- chromosome, as discussed in a later section. Most importantly, as shown in bold, tetraploid cells gave a significantly higher frequency of Neo+APRT- colonies, when both plasmids were present, than did diploid cells transfected with both plasmids or tetraploid cells transfected with pSXNeoTTAGGG alone.

The data in Table 1 suggested that a plasmid carrying a telomere sequence can join to an induced double-strand break and seed a new telomere. To examine that possibility further, we carried out a molecular analysis of 15 Neo+APRT- colonies that arose. In principle, Neo+APRT- colonies could arise in three ways: 1) independent integration and mutation, 2) integration at APRT, and 3) plasmid-to-chromsome-end joining followed by seeding. To distinguish these possibilities, we cut genomic DNA from individual colonies with BclI, which cleaves outside the APRT locus. The Southern blot of this digest was then successively hybridized to the Neo coding segment, the 3′ end of APRT, and the 5′ end of APRT. If integration and mutation were independent events, the neo and APRT probes would hybridize to different fragments; 8/15 colonies gave this pattern. If the neo plasmid integrated at the ARPT locus, both APRT probes and the neo probe would hybridize to the same fragment; 5/15 colonies gave this pattern. If the neo plasmid end-joined to one half of ARPT at the break and seeded a new telomere, the neo probe and one APRT probe would hybridize to the same fragment; 2/15 colonies gave this pattern.

The two colonies that show evidence for telomere seeding at the site of the induced double-strand break provided proof of principle that a seeded telomere at the APRT locus can heal a chromosome break, which was a fundamental assumption of our proposed research.

Orientation of the APRT gene on the CHO chromosome.

Before we began these experiments, the orientation of the APRT gene on the CHO chromosome was unknown. This is a critical piece of information, whose importance became relevant when the problem of chromosome loss was identified below. The experiments described above provided clear evidence for the orientation of the APRT gene on the chromosome. The structures of the two colonies that showed evidence for telomere seeding resolved the issue of the orientation of the APRT gene on the CHO chromosome. In both colonies the Neo gene was linked to the 3′ half of the APRT gene, while the 5′ half of the APRT gene was absent. Thus, the APRT gene is oriented with its 5′ end toward the telomere and the 3′ end toward the centromere. A defined orientation provided specific predictions for the outcomes of the experiments described in the next section and, in addition, provided a basis for our redesign of the test chromosome, as discussed later.

Double-strand breaks adjacent to interstitial telomere sequence and telomere seeding.

To determine whether a double strand break adjacent to the seeding end of an interstitial telomere sequence could promote chromosome truncation and the seeding of a new telomere, we first inserted an I-SceI recognition sequence adjacent to the seeding end of telomere sequences present in either orientation

in APRT gene targeting vectors. Targeting with these vectors created a duplication of the APRT gene, with the telomere sequence located in the second intron of the downstream APRT+ copy of the gene. Between the two copies of APRT sequence were located two other selectable markers—the herpesvirus TK and bacterial GPT gene—that have proven useful in the past (20-22). We created both orientations of the telomere sequence initially because we did not know the orientation of the APRT gene on the chromosome. The results described above gave the clear prediction that the CCCTAA orientation would permit truncation and seeding, while the TTAGGG orientation would not, serving as a control.

Tetraploid cell lines were created, as described above, for these two cell lines and for two other control cell lines: one that carried just the I-SceI site and one that carried just the telomere sequence in the CCCTAA orientation. Each of these cell lines was transfected with an I-SceI expression vector and then subjected to selection for APRT- cells and for APRT-TK- cells. The results of these experiments are shown in Table 2. We had anticipated that the I-SceI/CCCTAA cell line would show a substantial stimulation of APRT- and APRT-TK- colonies when transfected with the I-SceI expression vector. Although a slight stimulation was observed with both selections, the stimulation was not significantly different than that observed with the control, TTAGGG/I-SceI cell line. These data clearly identified the problem of chromosome loss.

Table 2. Effects of I-SceI expression in cell lines carrying an I-SceI recognition site adjacent to the seeding end of an ITS.

	APRT-			APRT-TK-		
Insert	**- I-SceI**	**+ I-SceI**	**Stimulat.**	**- I-SceI**	**+ I-SceI**	**Stimulat.**
I-SceI site	$3x10^{-4}$	$19x10^{-4}$	6	NA	NA	NA
CCCTAA	$11x10^{-4}$	$8x10^{-4}$	0.7	$3x10^{-4}$	$2x10^{-4}$	0.7
I-SceI/CCCTAA	$7x10^{-4}$	$10x10^{-4}$	1.4	$1.5x10^{-4}$	$3.6x10^{-4}$	2.4
TTAGGG/I-SceI	$14x10^{-4}$	$21x10^{-4}$	1.5	$8x10^{-4}$	$14x10^{-4}$	1.7

Chromosome loss.

As shown in Table 2, the background levels of APRT- and APRT-TK- cells obscured any possible effects of chromosome truncation and seeding. In the absence of I-SceI treatment all tetraploid cell lines had frequencies of APRT- cells on the order of 10^{-3}. Since the APRT locus is stable, we presumed that these high levels represent the frequency of chromosome loss in the tetraploid cells.

A high rate of chromosome loss interferes with our analysis as we had initially proposed to carry it out. Although we know we can truncate the APRT chromosome and seed a new telomere from the work described with plasmid vectors, we were not able to demonstrate it in the experimental paradigm we had proposed. In principle, we could avoid the problems of chromosome loss by anchoring the chromosome with a positively selectable marker, such as the GPT gene in our constructs. In our initial constructs, however, which were built before we knew the orientation of the APRT gene, the GPT gene was positioned so that it was lost when the chromosome was truncated. This necessitated a redesign of the test chromosome.

Redesign of the test chromosome.

Our original test chromosome was derived from a design that was appropriate for detecting homologous and nonhomologous recombination events at the APRT locus (see, for example, Appendix 1), which was the initial focus of our work on telomere sequence. The defined the orientation of the APRT locus on the chromosome allows us to explicitly represent the test chromosome as shown below.

Telomere---mutant 5′APRT---FRT---mutant 3′APRT---plasmid----TK----GPT---5′APRT---I-SceI/$(CCCTAA)_n$---FRT---3′APRT---Centromere

This test chromosome was generated by FLP/FRT site-specific recombination (23) between the circular targeting vector and the recipient chromosome as shown below. This site-specific recombination event could be selected for because it generated an APRT+ cell.

TARGETING VECTOR ---plasmid---TK---GPT---5′APRT---I-SceI/$(CCCTAA)_n$---FRT---mutant 3′APRT---

X

RECIPIENT CHROMOSOME Telomere---mutant 5′APRT---FRT---3′APRT----Centromere

The resulting test chromosome carried a doubly mutant APRT gene nearer the telomere and a functional APRT gene nearer the centromere. Both these genes carried an FRT site in the middle of their second intron; the FRT site in this location does not interfere with gene expression (21-23). The functional, centromere-proximal APRT gene additionally carries the I-SceI recognition site adjacent to the seeding end of 800 base pairs of telomere sequence, also in the second intron; these sequences do not interfere with gene expression. Between the duplicated copies of the APRT gene were located two other selectable marker genes: the herpesvirus thymidine kinase (TK) gene and the bacterial guanine phosphoribosyl transferase (GPT) gene. These two genes have proven useful in various selections for recombinants (20-22).

To detect chromosome truncation events, we prepared tetraploid cells to provide cover for any essential genes that might be located between the APRT locus and the telomere. This coverage was necessary because the APRT locus and surrounding chromosomal sequences are present in a single copy (haploid) in these CHO cells (24). Fusion was accomplished with a cell line that carried a defined deletion of the APRT locus (19). We anticipated that expression of I-SceI in these tetraploid lines would generate APRT– cells by chromosome truncation events. Detection of such events was obscured by the unexpectedly high frequency of chromosome loss, an unanticipated outcome.

These results required a redesign of the test chromosome so that chromosome loss (a frequency of 10^{-3}) would not interfere with our ability to detect chromosome truncation events. In addition we wanted to eliminate repeated APRT sequences to prevent homologous recombination (a frequency of about 10^{-4}). Eliminating homologous recombination and chromosome loss would lower the background events (derived from mutation and nonhomologous recombination) to less than 10^{-5}. The design of the new test chromosome is shown below.

TEST CHROMOSOME Telomere---FRT---plasmid---TK---I-SceI/$(CCCTAA)_n$---5′APRT---GPT---FRT---3′APRT---Centromere

This test chromosome has the desired features. It carries no duplicated APRT sequences so that the potential for homologous recombination is eliminated. A chromosome truncation event can be selected for as a TK–GPT+ cell. Chromosome loss will not be detected because it would generate TK–GPT– cells. In addition, the test chromosome could be used without further modification for transfer into breast cancer cells that have been rendered mutant for hypoxanthine-guanine phosphoribosyl transferase (HPRT), which is a straightforward procedure. Construction of this test chromosome required the targeting vector and recipient chromosome shown below.

TARGETING VECTOR ---plasmid---TK---I-SceI/(CCCTAA)n---5′APRT---GPT---FRT---

X

RECIPIENT CHROMOSOME Telomere---FRT---3′APRT---Centromere

In order to create the desired test chromosome, we needed to construct an appropriate, new targeting vector and recipient chromosome.

Construction of the recipient chromosome.

The recipient chromosome was constructed in several steps. First, a targeting vector that contained a precise deletion of the 5′APRT sequences (but retained sequences upstream of the APRT locus) was generated using appropriate flanking restriction sites. Next, it was introduced into a modified, functional APRT locus by site-specific recombination to generate an APRT– modified chromosome, as shown below.

TARGETING VECTOR ---plasmid---TK---upstream APRT---FRT---mutant 3′APRT—

X

CHROMOSOME Telomere---upstream APRT---5′APRT---FRT---3′APRT---Centromere

MODIFIED CHROMOSOME

Telomere---upstream APRT---5′APRT---FRT---mutant 3′APRT---plasmid---TK---upstream APRT---FRT---3′APRT---Centromere

Finally, the desired recipient chromosome was identified among TK– cells, some of which had undergone homologous recombination between the duplicated upstream APRT sequences to give the appropriate chromosome structure.

RECIPIENT CHROMOSOME Telomere---upstream APRT---FRT---3′APRT---Centromere

(Note that this recipient chromosome is identical to the one shown in the previous section. In this diagram, the upstream APRT sequences are explicitly shown because they are relevant to how the chromosome was generated. In the diagram in the previous section, the upstream APRT sequences are present, but not shown for simplicity because they are irrelevant to the proposed experiments.)

Construction of the targeting vector.

Repositioning the selectable markers and eliminating sequences from the APRT locus other than the 5′ half of the APRT gene required that we start over from scratch; we could not simply modify any existing targeting vector. In our design for the test chromosome, we wanted to move the selectable GPT gene as far away from the site of the truncation to insulate it from potential silencing effects caused by proximity to the new telomere created by truncation. The most distant location compatible with delivery by the targeting vector is adjacent to the FRT in the second intron of the APRT gene, as shown above. This design, however, required that both the APRT gene and the GPT gene can function when they were arranged as overlapping transcription units. To test this arrangement we cloned the GPT gene, in both orientations, adjacent to the FRT site in an otherwise wild-type APRT gene. These constructs were transfected into APRT– cells and parallel cultures were selected for either APRT+ colonies or GPT+ colonies. If there was any problem with overlapping transcription units, then APRT+ colonies, or GPT+ colonies, or both should be low. The results of these transfections revealed that both genes are functional regardless of the orientation of the GPT gene: APRT+ colonies and GPT+ colonies were obtained in equal numbers at high levels.

Having verified this element of the proposed targeting vector, we finalized a strategy for its construction. We designed a polylinker into which the various elements of the targeting vector could be cloned in a particular order so that there was minimum interference of one cloning step with the next. The resulting 70-nucleotide polylinker was constructed from overlapping shorter segments, inserted into a convenient plasmid backbone, and sequenced to verify its structure. The TK gene, the I-

SceI/$(CCCTAA)_n$ fragment, and the 5′APRT—GPT—FRT fragment were isolated from other vectors and inserted one-at-a-time into sites in the polylinker. The structure of the final targeting vector was verified by restriction digestion and Southern blotting.

Construction of a cell line carrying the redesigned test chromosome.

The targeting vector was transfected into the recipient cell line and APRT+ colonies were selected. Colonies were grown up and analyzed by Southern blotting to find colonies with the appropriately modified genome, containing the redesigned test chromosome. One such colony was subjected to the process of single-colony isolation to remove any contaminating parental cells, which are not uncommon in this type of selection (20). Isolated single colonies were grown and DNA was prepared and analyzed again by Southern blotting to confirm the structure of the test chromosome.

APRT+ tetraploid cell lines were created as described above by fusing this cell line with a cell line that was APRT- and ouabain resistant (18,19). The tetraploid character of the fused cells was verified by FACS analysis. When these tetraploid cells were exposed to dual selection for TK- and GPT+, they gave rise to the selected cells at a frequency of around 10^{-5}, as expected from our previous measures of the frequency of mutation of the TK gene in CHO cells. Thus, the redesign of the test chromosome appears to have had its intended effect: elimination of the problem of chromosome loss. We presume that chromosome loss still occurs, but such events no longer are recovered among our selected colonies.

We are currently exposing these cells to I-SceI endonuclease to see whether we get the expected increase in selected colonies, which would indicate that chromosome truncation along with the seeding of new telomeres can be specifically detected by selection for TK-GPT+ colonies. We will need to verify that the selected colonies do indeed possess the expected structure. If the redesigned test chromosome behaves as we expect it to, we will be in a position to carry out the experiments originally proposed in the grant application. With such a verified test chromosome in hand, we will undoubtedly be able to obtain funds to pursue aspects of the original proposal as well as others that have become relevant in the interim.

Construction of the color-based detection system for chromatid fusion events.

Our design for a color-based detection system depended on using the Green Fluorescent Protein (GFP). Initially, we inserted an artificial intron into an appropriate site in the GFP gene. Tests of that construct for ability to express GFP, which would require correct splicing of the added intron, were negative; only very low-level expression of GFP was obtained. Thus, it seems that consensus splicing signals are not all that is required for accurate splicing. Rather than work through the problem, as we have done in other studies to create artificial exons, we have used an alternative approach: fusion of GFP to a gene that already has introns. Such a fusion gene could be used in the way that was originally proposed. By inverting the initial exons we would eliminate GFP expression, and a chromatid fusion event adjacent to such a gene would re-link the initial exons from one chromatid with the distal exons from the other chromatid, thereby turning on expression of GFP. We have used for these purposes a C-terminal fusion of the human genomic rhodopsin gene with GFP, which we constructed for other purposes (25). The fused gene is under control of the CMV promoter and upon transfection into cells (or from a chromosomal site of integration) it expresses brightly green. We intend to use this available well-characterized rhodopsin-GFP fusion to serve as our color-based detector of chromatid fusion. Note that we have designed the polylinker in the redesigned targeting vector to have restriction sites where the rhodopsin-GFP fusion can be inserted in order to generate an appropriate test chromosome to follow

chromatid fusion. When the properties of the cell line with the redesigned test chromosome have been verified, we will insert the rhodopsin-GFP fusion gene in the form designed to detect chromatid fusion.

KEY RESEARCH ACCOMPLISHMENTS

- Quantification of the instability of interstitial telomere sequence
- Demonstration that a chromosomal double-strand break can serve as a site for plasmid capture and the seeding of a new telomere
- Determination of the orientation of the APRT gene on the CHO chromosome
- Identification of chromosome loss as a major confounding factor for the original experimental design
- Redesign of the test chromosome to avoid the problem of chromosome loss
- Construction of a new targeting vector for generating new test chromosome
- Construction of a recipient cell line for generating new test chromosome
- Construction of a cell line containing the newly designed test chromosome
- Verification that chromosome loss is not an obscuring factor in this new cell line
- Development a fusion of a genomic gene to the gene for GFP to serve as a color-based detector of chromatid fusion

REPORTABLE OUTCOMES

1. Manuscript: Kilburn, A.E., Shea, M.J., Sargent, R.G., and Wilson, J.H. "Insertion of telomere repeat sequence into a mammalian gene causes chromosome instability" *Mol. & Cell. Biol.* (submitted). Attached as Appendix 1.
2. Poster Abstract: Kilburn, A.E., Shea, M.J., Sargent, R.G., Sandor, Z., and Wilson, J.H. "Effects of telomere sequence on the stability and expression of a defined gene in mammalian cells" Prepared for the DOD Breast Cancer Research Program *Era of Hope*, June 8-12, 2000, in Atlanta, Georgia. Attached as Appendix 2.
3. Funding: An NIH grant is being prepared that contains the results described in this report and proposes to extend those results.
4. Degrees supported by this grant: It is anticipated that Jie Song, who was instrumental in construction of the new targeting vector (but was not paid from this grant), will obtain his Masters Degree within the next few months.

CONCLUSIONS

During the tenure of this grant, we have made progress toward our goal of delineating the initial steps in the process of telomere shortening and resultant cycles of fusion-bridge-breakage that lead to chromosome rearrangements. Most of the work we have accomplished has defined parameters of the test system we originally proposed to construct and use. We have shown that ITS at the APRT gene is some

30 fold more unstable than normal, but not so unstable as to prevent construction of the proposed test system. We have demonstrated that a plasmid bearing a telomere sequences can be captured at a DSB in the APRT gene and seed a new telomere, a proof in principle that our proposed experimental approach is viable. We identified a substantial, unanticipated difficulty in extending these studies to the chromosome: the problem of chromosome loss, which prevented us from observing chromosome truncation and telomere seeding at I-SceI-induced DSBs adjacent to an ITS. Given knowledge of the orientation of the APRT gene on the chromosome, which was defined in the plasmid-capture experiments, we were able to rationally design a new test chromosome that could be selectively anchored against chromosome loss. Engineering the redesigned test chromosome into CHO cells required a fundamental restructuring of the APRT system. We constructed new targeting vectors from scratch and engineered a recipient cell line that when joined to the targeting vector by site-specific FLP/FRT recombination would create the test chromosome. The new system has been constructed and tested. It avoids the problems of chromosome loss, which will permit us now to pursue the originally proposed experiments.

Beyond setting up a viable system to test the role of telomeres and telomere loss in the genomic instability associated with human breast cancer, we have defined the telomere repeat sequence as a destabilizing element in the interior of chromosomes in mammalian cells. The inherent instability of telomere sequence likely contributes to the rearrangements observed in cancer cells subsequent to the chromosome fusions that occur when telomeres become critically short (26-28). The extraordinarily high instability (several percent) correlated with some naturally occurring ITSs suggests that instability increases with telomere sequence length or with some undefined aspect of the arrangement or structure of the repeats (9-17). Our system also provides a means to quantify these undefined elements of telomere-sequence-induced instability.

REFERENCES

1. de Lange, T. 1995. Telomere dynamics and genome instability in human cancer. In Telomeres, Cold Spring Harbor Laboratory Press, New York, pp. 265-293.
2. Hastie, N.D. and R.C. Allshire. 1989. Human telomeres: fusion and interstitial sites. Trends Genet. **5**:326-331.
3. Fitzgerald, P.H. and C.M. Morris. 1984. Telomeric association of chromosomes in B-cell lymphoid leukemia. Hum. Genet. **67**:385-390.
4. McClintock, B. 1941. The stability of broken ends of chromosomes in *zea mays.* Genetics **26**:234-282.
5. McClintock, B. 1942. The fusion of broken ends of chromosomes following nuclear fusion. Proc. Natl. Acad. Sci. USA **28**:458-463.
6. Cowell, J.K. and O.J. Miller. 1983. Occurrence and evolution of homogeneously staining regions may be due to breakage-fusion-bridge cycles following telomere loss. Chromosoma **88**:216-221.
7. Sawyer, J., G. Sammartino, M. Husian, J.M. Lewis, B. Anderson, and F. Boop. 1993. Ring chromosome 12 resulting from nonrandom telomeric association with the short arm of chromosome 15 in cerebellar astrocytoma. Genes Chromosomes Cancer 8:69-73.
8. Blackburn, E.H., and C.W. Greider (eds.) 1995. Telomeres. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
9. Bouffler, S.D. 1998. Involvement of telomeric sequences in chromosomal aberrations. Mutat. Res. 404:199-204.
10. Boutouil, M., R. Fetni, J. Qu, L. Dallaire, C.L. Richer, and N. Lemieux. 1996. Fragile site and interstitial telomere repeat sequences at the fusion point of a de novo (Y;13) translocation. Hum. Genet. 98:323-327.
11. Cuthbert, G., S. McCullough, R. Finney, G. Breese, and N. Brown. 1999. Jumping translocation at 11q23 with MLL gene rearrangement and interstitial telomere sequences. Genes Chromosomes Cancer 24:295-298.

12. Day, J.P., C.L. Limoli, and W.F. Morgan. 1998. Recombination involving interstitial telomere repeat-like sequences promotes chromosomal instability in Chinese hamster cells. Carcinogenesis 19:259-265.
13. Jewett, T., D. Marnane, W. Stewart, R. Hayworth-Hodge, L. Finklea, K. Klinepeter, P.N. Rao, and M.J. Pettenati. 1998. Jumping translocation with partial duplications and triplications of chromosomes 7 and 15. Clin. Genet. 53:415-420.
14. Musio, A., G. Rainaldi, and I. Sbrana. 1996. Spontaneous and aphidicolin-sensitive fragile site 3cen co-localizes with the (TTAGGG)n telomeric sequence in Chinese hamster cells. Cytogenet. Cell Genet. 75:159-163.
15. Park, V.M., T.M. Gustashaw, and T.M. Wathen. 1992. The presence of interstitial telomeric sequences in constitutional chromosome abnormalities. Am. J. Hum. Genet. 50:914-923.
16. Slijepcevic, P., Y. Xiao, A.T. Natarajan, and P.E. Bryant. 1997. Instability of CHO chromosomes carrying interstitial telomere sequences originating from Chinese hamster chromosome 10. Cytogenet. Cell Genet. 76:58-60.
17. Vermeesch, J.R., P. Petit, F. Speleman, K. Deuriendt, J.P. Fryns, and P. Marynen. 1997. Interstitial telomeric sequences at the junction site of a jumping translocation. Hum. Genet. 99:735-737.
18. Adair, G.M., L.H. Thompson, and P.A. Lindl. 1978. Six complementation classes of conditionally lethal protein synthesis mutants of CHO cells selected by 3H-amino acid. Somat. Cell Genet. 4:27-44.
19. Nairn, R.S., R.M. Humphrey, and G.M. Adair. 1988. Transformation of UV-hypersensitive Chinese hamster ovary cell mutants with UV-irradiated plasmids. Int. J. Radiat. Biol. 53:249-260.
20. Sargent, R.G., R.V. Merrihew, R. Nairn, G. Adair, M. Meuth, and J.H. Wilson. 1996. The influence of a (GT)29 microsatellite sequence on homologous recombination in the hamster APRT gene. Nucl. Acids Res. 24:746-753.
21. Sargent, R.G., M. Brenneman, and J.H. Wilson. 1997. Repair of site-specific double-strand breaks on a mammalian chromosome by homologous and illegitimate recombination. Mol. and Cell Biol. 17:267-277.
22. Sargent, R.G., R.L. Rolig, A.E. Kilburn, G.M. Adair, J.H. Wilson, and R.S. Nairn. 1997. Recombination-dependent deletion formation in mammalian cells deficient in the nucleotide excision repair gene ERCC1. Proc. Natl. Acad. Sci. USA 94:13122-13127.
23. Merrihew, R.V., R.G. Sargent, and J.H. Wilson. 1995. Efficient modification of the APRT gene by FLP/FRT site-specific targeting. Som. Cell & Mol. Genet. 21:299-307.
24. Adair, G.M., R.L. Stallings, R.S. Nairn, and M.J. Siciliano. 1983. High-frequency structural gene deletion as the basis for functional hemizygosity of the adenine phosphoribosyltransferase locus in Chinese hamster ovary cells. Proc. Natl. Acad. Sci. USA 80:5961-5964.
25. Intody, Z, B.D. Perkins, J.H. Wilson, and T.G. Wensel Blocking transcription of the human rhodopsin gene by triplex-mediated DNA photocrosslinking. Nucl. Acids Res. (in press).
26. Counter, C.M., H.W. Hirte, S. Bacchetti, and C.B. Harley. 1994. Telomerase activity in human ovarian carcinoma. Proc. Natl. Acad. Sci USA 91:2900-2904.
27. Harley, C.B., N.W. Kim, K.R. Prowse, S.L. Weinrich, K.S. Hirsch, M.D. West, S. Bacchetti, H.W. Hirte, C.W. Greider, W.E. Wright, and J.W. Shay. 1994. Telomerase, cell immortality and cancer. In Cold Spring Harbor Symposia in Quantitative Biology, Vol. LVIX, pp 307-315. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
28. Hastie, N.D., and R.C. Allshire. 1989. Human telomeres: fusion and interstitial sites. Trends Genet. 5:326-331.

APPENDICES:

1. Manuscript: Kilburn, A.E., Shea, M.J., Sargent, R.G., and Wilson, J.H. "Insertion of telomere repeat sequence into a mammalian gene causes chromosome instability" Mol. & Cell. Biol. (submitted).

2. Poster Abstract: Kilburn, A.E., Shea, M.J., Sargent, R.G., Sandor, Z., and Wilson, J.H. "Effects of telomere sequence on the stability and expression of a defined gene in mammalian cells" Prepared for the DOD Breast Cancer Research Program *Era of Hope*, June 8-12, 2000, in Atlanta, Georgia.

BIBLIOGRAPHY:

1. Kilburn, A.E., Shea, M.J., Sargent, R.G., and Wilson, J.H. "Insertion of telomere repeat sequence into a mammalian gene causes chromosome instability" Mol. & Cell. Biol. (submitted).
2. Kilburn, A.E., Shea, M.J., Sargent, R.G., Sandor, Z., and Wilson, J.H. "Effects of telomere sequence on the stability and expression of a defined gene in mammalian cells" Abstract prepared for the DOD Breast Cancer Research Program *Era of Hope*, June 8-12, 2000, in Atlanta, Georgia.

LIST OF PERSONNEL RECEIVING PAY FROM RESEARCH EFFORT:

John Wilson
Fung Chan

INSERTION OF TELOMERE REPEAT SEQUENCE INTO A MAMMALIAN GENE CAUSES CHROMOSOME INSTABILITY

April E. Kilburn*†‡, Martin J. Shea•, R. Geoffrey Sargent*†, and John H. Wilson*•†#

*The Verna and Marrs McLean Department of Biochemistry and Molecular Biology

•Department of Molecular and Human Genetics

†Cell and Molecular Biology Program

Baylor College of Medicine

Houston, TX 77030

Running Title: Telomere Repeat-Mediated Instability

† Present address: Pangene Corporation, Mountain View, CA 94043

‡ Present address: Office of Technology and Licensing, University of Texas, Austin, TX 78759

Corresponding author. Mailing address: Department of Biochemistry and Molecular Biology, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030

Phone: (713) 798-5760. Fax: (713) 796-9438. E-mail: jwilson@bcm.tmc.edu

ABSTRACT

Telomere repeat sequences cap the ends of eucaryotic chromosomes and help stabilize them. At interstitial sites, however, they may destabilize chromosomes, as suggested by cytogenetic studies in mammalian cells that correlate interstitial telomere sequence with sites of spontaneous and radiation-induced chromosome rearrangements. In no instance is the length, purity, or orientation of the telomere repeats at these potentially destabilizing interstitial sites known. To determine the effects of a defined interstitial telomere sequence on chromosome instability, as well as other aspects of DNA metabolism, we deposited 800 bp of the functional vertebrate telomere repeat, TTAGGG, in two orientations in the second intron of the adenosine phosphoribosyltransferase (*APRT*) gene in Chinese hamster ovary cells. In one orientation the deposited telomere sequence did not interfere with expression of the *APRT* gene, whereas in the other it reduced mRNA levels by about 50%. The telomere sequence did not induce chromosome truncation and the seeding of a new telomere at a frequency above the limits of detection. Similarly, the telomere sequence did not alter the rate or distribution of homologous recombination events. The interstitial telomere repeat sequence in both orientations, however, dramatically increased gene rearrangements some 30-fold. Analysis of individual rearrangements confirmed the involvement of the telomere sequence. These studies define the telomere repeat sequence as a destabilizing element in the interior of chromosomes in mammalian cells.

INTRODUCTION

Several kilobases of short repeated sequences—TTAGGG in vertebrates—make up the DNA component of telomeres, which cap the ends of eucaryotic chromosomes (9). These sequences serve as binding sties for a collection of proteins that compensate for progressive losses due to replication (9), protect the ends from nuclease degradation and end-to-end fusion (11, 70), and give rise to a unique chromatin structure (69). Telomeric proteins play additional roles in chromosome attachment to the nuclear matrix (44) and in the separation of telomeres at mitosis and meiosis (13, 40). Thus, the telomere sequence mediates a complicated interplay of proteins and processes.

Telomeres influence replication, gene expression, and recombination in their vicinity. Activation of replication origins is delayed or abolished near telomeres in mitotically dividing *S. cerivisiae* (21, 26, 56, 68). Genes near telomeres in *S. cerivisiae* (28), *S. pombe* (52), *Drosophila melanogaster* (43), and *Trypanosome brucei* (33, 60) are transcriptionally repressed. Near telomeres in mammalian cells, selectable genes with strong promoters are not significantly affected, whereas genes driven by weak promoters can be repressed (8, 14). During meiosis in *S. cerevisiae*, ectopic recombination is significantly greater between inserts near telomeres than it is between more centrally located inserts (27), although recombination between directly repeated *LUE2* gene segments was unaffected by proximity to the telomere (54). In humans, meiotic recombination is elevated near telomeres (5, 39). By contrast, molecular and cytological studies of meiosis in grasshopper show reduced recombination near telomeres (48).

Telomere repeats are not confined to the ends of chromosomes, but are also found at discrete intrachromosomal sites in many eucaryotic species (1, 6, 19, 55). It is thought

that these interstitial telomere repeats arose as the result of chromosome rearrangements in the course of genome evolution (34, 66), a view supported by occasional observation of aberrant chromosomes that have telomere repeats at the site of rearrangement (57). Like repeats at telomeres, interstitial repeats also appear to influence aspects of DNA metabolism in their vicinity. Cytogenetic studies in mitotically dividing cells have linked interstitial telomere repeats with sites of spontaneous and radiation-induced chromosome rearrangements (10, 17, 53, 65), chromosome fragility (12, 50), and unstable rearrangements known as jumping translocations (16, 36, 71). In meiotic cells in the Armenian hamster an interstitial telomere repeat was a site of frequent chiasma formation, consistent with a hotspot for homologous recombination (4). DNA molecules injected into the macronucleus of *Paramecium primaurelia* preferentially integrate by illegitimate recombination in or near interstitial telomere repeats (37).

Because interstitial telomere sequences are uncharacterized for length, purity, and repeat orientation and because interstitial repeats are not all hotspots for rearrangement (10), several studies introduced defined telomeric sequences into the genome. In *S. cerevisiae*, insertion of 49 bp of telomeric sequence at the *HIS4* locus stimulated meiotic homologous recombination and the formation of nearby meiosis-specific double-strand DNA breaks (22, 73). In mitotic yeast cells, homologous recombination between 300 basepair duplications of telomeric sequence occurred at roughly the same frequency as that between the same length of unique sequence, except in the vicinity of the telomere where telomere repeat recombination was reduced tenfold (67). Overexpression of the telomere-binding protein Rap1p eliminated repression of recombination near telomeres and stimulated recombination at interior telomeric repeats, indicating that some telomere-

repeat binding proteins recognize interstitial sequences (67). Finally, at several locations in the *S. cerevisiae* genome, telomeric repeats repress transcription of nearby genes (67).

In mammalian cells telomere repeat sequences have been introduced to fragment chromosomes and to generate minichromosomes (7, 23, 24, 25, 29, 32, 35, 41, 49). Random integration of plasmids carrying telomere repeats adjacent to a selectable marker generated selected colonies with a newly seeded telomere next to the marker at a frequency of 20% in Chinese hamster ovary (CHO) cells (23) and 70% in HeLa cells (29). Surprisingly, the majority of such clones carried duplications or other rearrangements at the site of chromosome truncation (14, 24, 32). The role of telomere sequence in chromosome truncation and terminal rearrangement—beyond its capacity to seed new telomeres—is unclear. Cytogenetic analysis of three human cell lines with randomly integrated telomere-repeat-containing plasmids showed that two were highly unstable, but the instability was not due to telomere sequence (20), but rather to random integration, which commonly generates ongoing rearrangements (47, 58).

To assess the effects of interstitial telomere sequence on several aspects of DNA metabolism, we used site-specific recombination to insert 800 bp of functional vertebrate telomere sequence in two orientations into the second intron of the adenosine phosphoribosyltransferase (*APRT*) gene in CHO cells. Site-specific recombination avoids the inherent instability of many random integrants (20, 47, 58), and targeting to the *APRT* locus allows us to make comparisons with previous results (62, 63, 64). These cell lines allowed us to test the effects of telomere sequence on gene expression, homologous recombination, gene rearrangements, and chromosome truncation.

MATERIALS AND METHODS

Construction of vectors. Targeting vectors were constructed from previously described vectors (62, 64), which contained the herpesvirus *TK* gene, the bacterial *GPT* gene, and an *APRT*$^-$ gene truncated at the 3′ end of the last exon. To generate targeting vector pAK30 containing telomere sequence in the CCCTAA orientation, telomere sequence from the Sty11 plasmid, kindly provided by Dr. Titia de Lange, (29) was subcloned into a polylinker adjacent to an FRT site and the pair were then cloned into the polylinker in pGS89. To construct targeting vector pAK50 containing telomere in the TTAGGG orientation, telomere sequence was cloned into the polylinker adjacent to the FRT site in pGS101. Orientations of the telomere sequence are indicated by the sequence of the repeat in the mRNA-like strand of the DNA. Targeting vectors pAK30 and pAK50 were checked for the presence and correct orientation of telomere sequence by restriction digestion of surrounding polylinker sequence and by sequence analysis. In both cases the telomere sequence consisted predominantly of TTAGGG repeats, with interspersed TTGGGG repeats common at the nonseeding (TA-rich) end but rare at the seeding (G-rich) end. The telomere sequences in pAK30 and pAK50 were identical to that in Sty11, with the exception that both were missing one TTGGGG repeat at the nonseeding end. Targeting vectors containing the I-SceI recognition site were constructed by inserting a synthetic I-SceI site into a restriction site in the polylinker adjacent to the seeding end of the telomere sequence.

A targeting vector containing *HPRT* DNA was constructed by ligating an 800-bp PCR fragment (from base 14928 to base 15730 of the human *HPRT* intron 2) into the SalI and NotI sites in the polylinker at the EcoRI site in pGS101, via SalI and NotI sites in the

PCR primers. The targeting vector containing *HPRT* DNA with a central I-SceI site was constructed by recombinant PCR. The outside primers were the same as above and the inside primers created an I-SceI site.

Vectors for random integration were constructed by modification of plasmid pGS36, which carried the wild-type *APRT* gene with an adjacent *GPT* gene and 4.5 kb of upstream sequences. A HindIII-XhoI fragment from pAK30 or pAK50, which includes the upstream sequences, the *GPT* gene, and a segment of the *APRT* gene containing exons 1 through the middle of exon 3 and the telomere sequences, was used to replace the corresponding segment of pGS36 to generate plasmids pAK301 and pAK501. These vectors were linearized at the unique HindIII site prior to transfection.

Construction of cell lines. Site-specific recombination between the FRT sites in the vectors and in the endogenous *APRT* gene on the chromosome was carried out as described previously (46). The *APRT* gene in the RMP41 cell line (46) carries a nonreverting point mutation that eliminates the EcoRV site in exon 2 (62). Site-specific recombination generated the tandemly duplicated gene structures shown in Figure 1. The upstream *APRT* gene carries two mutations: the point mutation in exon 2 and the truncation of the 3′ end. The downstream, functional *APRT* gene carries the telomere sequence or *HPRT* DNA in intron 2. Cell line AK550 was derived from cell line AK213 by selection for $TK^{-}APRT^{+}$ colonies arising by homologous recombination (62).

Tetraploid cell lines were constructed by cell fusion (2) between the tandem duplication cell lines carrying telomere sequence and a ouabain-resistant derivative of T2S24 (51), in which exons 1 and 2 of the *APRT* gene are deleted. Fused cells were selected by growth in medium containing 1 mM ouabain and ALASA (25 μM alanosine,

50 μM azaserine, and 100 μM adenine) (2), and shown to be tetraploid by FACS analysis.

Structures of all cell lines were verified by Southern analysis following digestion with restriction enzymes diagnostic for the predicted structure.

The parental cell line and selected diploid and tetraploid cell lines were shown to contain active telomerase by telomeric repeat amplification protocol (TRAP) assays (38).

Cell culture, fluctuation analysis, and transfection. Cell lines were maintained in Dulbecco's modified Eagle medium supplemented with amino acids and 10% fetal calf serum. Selections were carried out as previously described (62). *APRT*$^+$ cells were selected by growth in ALASA medium. *APRT*$^-$ cells were selected by growth in medium made with 10% dialyzed fetal calf serum and supplemented with 400 μM 8-azaadenine. *TK*$^-$*APRT*$^-$ cells were selected by growth in *APRT*$^-$ selection medium supplemented with 0.3 μM fluoroiodoarabinosyluridine (FIAU).

Fluctuation analysis (42, 45) was carried out using 12 parallel cultures grown from initial populations of 50 to 100 cells for each rate determination, as described previously (62). The numbers of *APRT*$^-$ or *TK*$^-$*APRT*$^-$ colonies in parallel cultures were used to calculate rates by the method of the median (42). A single colony was picked from each parallel culture to ensure that all analyzed colonies arose independently.

In experiments that used I-SceI to generate double-strand breaks 15 μg of the expression vector for I-SceI, pCMVI-SceI (59), was introduced by LipofectAmine (Gibco/BRL) into subconfluent cultures on 100 mm plates, as described previously (63).

Southerns, Northerns, PCR analysis, and DNA Sequencing. Northern and Southern analyses were carried out using standard protocols (61). The probe for Southern analysis

was the 3.9-kb BamHI fragment containing the entire *APRT* gene, labeled by random priming with [^{32}P]dCTP. The probes for the Northern blot were a CHO *APRT* cDNA, kindly provided by Elliot Drobetsky, and GAPDH cDNA as an internal loading control. Quantification of RNA on Northern blots was performed by a Phosphorimager using Molecular Dynamics software. PCR analysis of the recombination products was carried out as previously described (62). The locations of PCR primers used for analysis of rearrangements are shown in Figure 4A; their sequences are available on request. DNA sequencing was carried out using automated sequencing technology on targeting plasmids to confirm the orientation of the telomere sequence insert and on amplified PCR fragments to determine the sequences of the rearrangement junctions.

RESULTS

Effects of telomere sequence on the *APRT*$^+$ phenotype. In preparation for the experiments described here, it was essential to know that telomere sequence in either orientation would not interfere with the ability of CHO cells to express the *APRT*$^+$ phenotype. A functional *APRT* gene was necessary for our targeting strategy and for our loss-of-function assays for homologous recombination, gene rearrangement, and chromosome truncation. To address this question, we linearized plasmid pGS36 (no insert), pAK301 (CCCTAA), and pAK501 (TTAGGG) and transfected them into the *APRT*$^-$ cell line RMP41. Transfected cells were plated to recover *APRT*$^+$ or *GPT*$^+$ colonies arising by random integration of the plasmid DNA (Table 1). If the telomere sequence embedded in the middle of the *APRT* gene blocked its expression, we would have expected many fewer *APRT*$^+$ colonies than *GPT*$^+$ colonies. Because *APRT*$^+$ and *GPT*$^+$ colonies were recovered at roughly equal frequencies in transfections with each plasmid, we concluded that 800 bp of telomere sequence in the second intron did not affect the ability of the gene to express the *APRT*$^+$ phenotype.

Construction of cell lines and experimental rationale. To test the effects of interstitial telomere sequence on gene expression, homologous recombination, gene rearrangements, and chromosome truncation, we constructed a variety of cell lines whose structures are shown in Figure 1A. Targeting vectors carrying different DNA sequences adjacent to an FRT site in the second *APRT* intron were integrated via FLP-mediated site-specific recombination so that the inserted sequences were located in the downstream, *APRT*$^+$ copy of the gene. The upstream *APRT*$^-$ copy of the gene carries a nonreverting point mutation and is truncated at its 3′ end. The structures of the tandem duplications in these

cells lines is analogous to those we have used before (62, 64) and thus allow us to make direct comparisons with our previous results.

We chose to use 800 bp of telomere sequence because this length was sufficient to support telomere-associated chromosome fragmentation (TACF) in HeLa cells (29). Similar experiments in CHO cells demonstrated TACF using only 500 bp of telomere sequence (23). As a control for this length of insert, we constructed cell lines carrying an 800-bp fragment from intron 2 of the human *HPRT* gene (Figure 1A). As a positive control to test the effects of double-strand breaks, we constructed parallel cell lines that carried the recognition site for the I-SceI endonuclease at the seeding (G-rich) end of the telomere sequence or in the middle of the *HPRT* fragment (arrows in Figure 1A).

To test for chromosome truncation accompanied by the formation of new telomeres at the inserted telomere sequence, it was necessary to provide a second copy of the chromosome that carries the *APRT* gene. Since large portions of the *APRT*-containing chromosome are hemizygous (3), it was possible that chromosome truncation would eliminate an essential gene, rendering those cells nonviable. A second copy of the *APRT*-containing chromosome was provided by fusion with cell line T2S24, which carries a deletion in the *APRT* gene (51). A defective *APRT* gene on this second chromosome was critical for our analysis, which depends on selection for the $APRT^-$ phenotype.

Effects of telomere sequence on production of *APRT* mRNA. Although telomere sequence does not interfere with expression of the $APRT^+$ phenotype (Table 1), it could still reduce mRNA levels substantially, since cells with only a few percent of wild-type Aprt enzyme activity are phenotypicaly $APRT^+$ (18). To measure the effect of telomere sequence on production of *APRT* mRNA, we performed Northern analysis on RNA

extracted from wild-type cells and from cells with targeted tandem duplication at the *APRT* locus (Figure 2). The level of *APRT* mRNA relative to GAPDH mRNA is the same for control cell lines carrying a single copy of the *APRT* gene (lane 1), a tandem duplication with no insert (lane 3), and a tandem duplication with the *HPRT* insert (lane 4). Cell line AK213 with telomere sequence in the CCCTAA orientation expresses the same relative level of *APRT* mRNA as the control cell lines (lane 5). Cell line AK775 with telomere sequence in the TTAGGG orientation expresses about half the amount of *APRT* mRNA, relative to GAPDH mRNA, as the other cell lines (lane 6).

Since we did not know the transcriptional status of the upstream *APRT* gene fragment, we included in the analysis an additional cell line, DELI26, which carries a tandem duplication with the same structure, except that the downstream gene is promoterless (46). The absence of *APRT* mRNA in this cell line (lane 2) indicates that the upstream fragment of the *APRT* gene does not yield a stable transcript. Thus, the detected transcripts in the other tandem-duplication cell lines must come exclusively from the downstream gene, which harbors the different sequences tested. From these experiments we conclude that the CCCTAA orientation of telomere sequence has no effect on *APRT* transcription and processing, whereas the TTAGGG orientation reduces mRNA production by about 50%.

Effects of telomere sequence on chromosome truncation. The single functional *APRT* gene in CHO AT32 cells resides in a hemizygous region near the chromosome end and is transcribed toward the centromere (72). Truncation of the chromosome at the interstitial telomere sequence and the seeding of a new telomere, which could occur only in the CCCTAA orientation, would eliminate the 5′ end of the *APRT* gene along with more

distal sequences. To render eliminated sequences nonessential, we fused tandemly duplicated cell lines to a cell line carrying an *APRT* deletion. In these tetraploid cell lines, *APRT*$^-$ colonies arose (presumably due to chromosome loss) at the same frequency in cell lines with no telomere sequence (AK723, $3.8 \pm 1.8 \times 10^{-4}$) and with a nonseeding, TTAGGG sequence (AK858, $4.9 \pm 1.9 \times 10^{-4}$). In two potentially seeding, CCCTAA cell lines (AK863, AK728) *APRT*$^-$ colonies arose at a similar frequency (average $6.6 \pm 3.1 \times 10^{-4}$). Thus, telomere sequence at the *APRT* locus does not cause chromosome truncation and the seeding of new telomeres at a frequency greater than 0.1%. Expression of I-SceI in these cell lines did not stimulate *APRT*$^-$ colony formation sufficiently above the level of chromosome loss to detect the seeding of new telomeres directly (data not shown).

Effects of telomere sequence on homologous recombination. Although tandem duplications can give rise to *APRT*$^-$ cells in several ways (Figure 1B), previous analysis of spontaneous events indicated that homologous recombination was dominant, accounting for about 95% of events as compared to 5% for mutations and <0.5% rearrangements (62, 64). The same studies showed that *TK*$^-$*APRT*$^-$ cells were generated entirely by homologous recombination (Figure 1B). Thus, we measured the effects of telomere sequence on homologous recombination by measuring rates of production of *APRT*$^-$ and *TK*$^-$*APRT*$^-$ phenotypes. Homologous recombination yields *TK*$^-$*APRT*$^-$ cells by crossover (popout) recombination, which eliminates one copy of the *APRT* gene; it generates *APRT*$^-$ cells by crossover recombination and by gene conversion, in which the EcoRV mutation in the upstream copy is transferred to the downstream copy (Figure 1B).

Tandem duplications carrying telomere sequence yielded *TK*$^-$*APRT*$^-$ cells and *APRT*$^-$ cells at rates that were indistinguishable from those carrying an *HPRT* insert or

smaller inserts (Table 2), suggesting that homologous recombination was unaffected by telomere sequence in either orientation. Analysis of individual colonies by Southern blotting and PCR confirmed that the majority arose by homologous recombination (Figure 3, Table 2). Among *APRT*$^-$ colonies from cell lines containing telomere sequence, the proportions of conversions and crossovers (18 *vs* 4) was similar to that observed previously (88 *vs* 17) (Table 2). In addition, crossovers that retained telomere sequence (13/24) or lost it (11/24) were generated in proportion to the lengths of homology flanking the telomere sequence, consistent with results for smaller inserts (62). Thus, neither the rates of recombination nor the nature of the products reveals any influence of telomere sequence.

To determine whether stimulated recombination could be detected in the vicinity of telomere sequence, we expressed I-SceI endonuclease in cell lines that carried its recognition site alone, adjacent to the telomere sequence, or in the middle of the *HPRT* insert (Figure 1A). I-SceI expression typically stimulates homologous recombination several hundred fold in mammalian cells (59, 63). In all cases recombination was stimulated to similar levels (Table 3), much above the spontaneous rates (Table 2). Analysis of individual colonies confirmed that most arose by homologous recombination (data not shown). Since increased recombination could have been detected, we conclude that telomere sequence does not affect homologous recombination at *APRT*.

Effects of telomere sequence on gene rearrangements. Our loss-of-function assay allows us to detect gene rearrangements in addition to homologous recombinants (Figure 1B). Rearrangements had not previously been observed among spontaneous recombinants in wild-type CHO cells (62, 64). Thus, the most striking feature of the data

in Table 2 is the presence of rearrangements, which were detected by their abnormal Southern patterns (Figure 3). Among 59 colonies from tandem duplications carrying the telomere sequence, 9 were rearrangements. By contrast, 0 of 23 colonies from the *HPRT* insert were rearrangements. In experiments with inserts of less than 200 bp (62, 64), no rearrangements were detected among 183 analyzed colonies (Table 2). These numbers (9/59 for telomere sequence versus 0/206 for *HPRT* and small inserts) indicate that rearrangements were stimulated some 30-fold or more by interstitial telomere sequence. One additional rearrangement was found among ten independent *APRT*$^-$ colonies isolated from cell line AK550, which had a single copy of the *APRT* gene (Figure 1A).

PCR and Southern analysis of the rearrangements showed that they included deletions, insertions, and a probable translocation and that all involved the telomere sequence (Figure 4A). Five rearrangement junctions were successfully amplified by PCR and sequenced (Figure 4B). In rearrangement AK586 *APRT* sequences around the original telomere sequence were deleted and replaced with a short telomere sequence, flipped with respect to the original orientation. The new telomere insert does not correspond to a known sequence in the original insert, and thus it may have been corrupted in the course of the rearrangement.

Rearrangements AK797, AK802, AK806, and AK810 each had lost some 400 bp of telomere sequence. Interspersed TTGGGG repeats at the nonseeding end of the telomere sequence allowed one end of the deletion to be mapped to a common region in all four rearrangements (Figure 4B); however, their 3′ ends could not be positioned relative to the featureless repeats at the seeding end of the telomere sequence. Differences in lengths of Southern fragments and PCR products suggest that these

telomere-sequence deletions are not identical. Nevertheless, their similarity raised the possibility that they pre-existed in the starting cell population. If their isolation in these experiments was coincidental, then they should be present in the starting population at roughly the same frequency that they appeared among the selected colonies; that is, at 15 to 20 % (4 isolates among 22 colonies). Sensitive Southern analysis, which could have detected about a 2% subpopulation, failed to reveal the band diagnostic for these telomere-sequence deletions in the genomic DNA from the parental AK775 cell line (data not shown).

DISCUSSION

These studies clearly document that interstitial telomere repeat sequence of known purity, specific length, and in either orientation confers instability at a defined site in a mammalian genome. By using the well-characterized *APRT* locus in CHO cells and a substrate design that allowed sensitive, simultaneous detection of homologous recombination and gene rearrangements, we have demonstrated that telomere sequence stimulates rearrangements some 30-fold above background, without noticeable effects on homologous recombination. Direct participation of telomere sequence in the detected rearrangements is supported by molecular characterization, which showed that every rearrangement involved the telomere repeats. Previous studies using random integration of a 1.6-kb telomere sequence failed to detect repeat-induced instability by less-sensitive cytogenetic methods (20). Thus, these studies define the telomere repeat sequence as a destabilizing element in the interior of a mammalian chromosome, providing direct support for previous correlations between interstitial telomere repeats and chromosome rearrangements (10, 12, 16, 17, 36, 50, 53, 65, 71).

The effects of telomeres on expression of nearby genes is dramatic in many lower eucaryotes (28, 33, 43, 52, 60), but relatively weak in mammalian cells (8, 14). In yeast interstitial telomere sequence also reduces expression of nearby genes (67). In our studies we have shown that telomere sequence in an intron of the *APRT* gene has only a modest effect on expression of the gene (Figure 2). Because only one orientation of the repeat (TTAGGG) reduced mRNA levels, it seems unlikely that the reduction is due to repeat-binding proteins, whose effects might be expected to be orientation independent. It may be that the repeated sequence in the template strand (3′-AATCCC) impedes RNA

polymerase, or that the repeated sequence in the nascent RNA (5′-UUAGGG) interferes with RNA processing. Further studies would be required to resolve these possibilities.

Chromosome truncation and the seeding of new telomeres were not detected above the background loss of chromosomes that is common in tetraploid cell lines (2), which places an upper limit of about 0.1% on the frequency of these events at the *APRT* locus. When these studies were initiated, the orientation of the *APRT* gene on the chromosome was unknown (72) and our particular arrangement of selectable markers did not allow us to distinguish between a lost chromosome and a truncated one. We have now reconfigured the markers to address this issue with more sensitivity. Nevertheless, the low frequency of chromosome truncation cannot account for the 20-70% truncation frequencies observed in TACF experiments (23, 29), suggesting that TACF is unlikely to occur by random integration followed by telomere-sequence-induced breakage. It seems more likely that truncation observed in TACF experiments results from plasmid ligation to transient double-strand breaks, or from random-integration-triggered rearrangements that are resolved when a break appears near the telomere sequence (14).

The effects of telomeres and telomere sequence on homologous recombination are varied, sometimes stimulating it (4, 5, 22, 27, 39, 73), sometimes inhibiting it (48, 67), and sometimes leaving it unaffected (54). At the *APRT* locus telomere sequence does not detectably affect homologous recombination, as assessed by rates of recombination, proportions of crossovers and conversions, and distribution of exchanges. When double-strand breaks were deliberately introduced adjacent to telomere sequence, homologous recombination was stimulated to the same extent as in cell lines lacking telomere

sequence. Thus, stimulated recombination could have been detected in the vicinity of the telomere sequence.

The inherent instability of interstitial telomere sequence likely contributes to the rearrangements observed in cancer cells subsequent to the chromosome fusions that occur when telomeres become critically short (15, 30, 31). The extraordinarily high instability (several percent) correlated with some naturally occurring interstitial telomere sequences (10, 12, 16, 17, 36, 50, 53, 65, 71), suggests that instability increases with telomere sequence length or with some undefined aspect of the arrangement or structure of the repeats. The approaches described here provide a means to quantify these undefined elements of telomere-sequence-induced instability.

ACKNOWLEDGEMENTS

We wish to thank Drs. Gerald Adair and Olivia Perrera-Smith for advice about cell fusions, Dr. Dan Medina for help with telomerase assays, and Drs. Beth and Frank Chance for helpful discussions. This investigation was supported by NIH grant GM38219 and by Department of Defense Breast Cancer Research Program grant DAMD17-97-1-7283 to J.H.W.

References:

1. Abuin, M., P. Martinez, and L. Sanchez. 1996. Localization of the repetitive telomeric sequence (TTAGGG)n in four salmonoid species. Genome **39**:1035-1038.
2. Adair, G.M., L.H. Thompson, and P.A. Lindl. 1978. Six complementation classes of conditionally lethal protein synthesis mutants of CHO cells selected by ^{3}H-amino acid. Somat. Cell Genet. **4**:27-44.
3. Adair, G.M., R.L. Stallings, R.S. Nairn, and M.J. Siciliano. 1983. High-frequency structural gene deletion as the basis for functional hemizygosity of the adenine phosphoribosyltransferase locus in Chinese hamster ovary cells. Proc. Natl. Acad. Sci. USA **80**:5961-5964.
4. Ashley, T., and D.C. Ward. 1993. A "hotspot" of recombination coincides with an interstitial telomeric sequence in the Armenian hamster. Cytogenet. Cell Genet. **62**:169-171.
5. Ashley, T. 1994. Mammalian meiotic recombination: A reexamination. Hum. Genet. **94**:587-593.
6. Azzalin, C.M., E. Mucciolo, L. Bertoni, and E. Giulotto. 1997. Flourescence in situ hybridization with a synthetic (TTAGGG)n polynucleotide detects several intrachromosomal telomere-like repeats on human chromosomes. Cytogenet. Cell. Genet. **78**:112-115.
7. Barnett, M.A., V.J. Buckle, E.P. Evans, A.C. Porter, D. Rout, A.G. Smith, and W.R. Brown. 1993. Telomere directed fragmentation of mammalian chromosomes. Nucl. Acids Res. **21**:27-36.

8. Bayne, R.A.L., D. Broccoli, M.H. Taggart, E.J. Thomson, C.J. Farr, and H.J. Cooke. 1993. Sandwiching of a gene within 12 kb of a functional telomere and alpha satellite does not result in silencing. Hum. Genet. **3**:539-546.

9. Blackburn, E.H., and C.W. Greider (eds.) 1995. Telomeres. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

10. Bouffler, S.D. 1998. Involvement of telomeric sequences in chromosomal aberrations. Mutat. Res. **404**:199-204.

11. Bourgain, F.M., and M.D. Katinka. 1991. Telomeres inhibit end to end fusion and enhance maintenance of linear DNA molecules injected into the *Paramecium primaurelia* macronucleus. Nucl. Acids Res. **19**:1541-1547.

12. Boutouil, M., R. Fetni, J. Qu, L. Dallaire, C.L. Richer, and N. Lemieux. 1996. Fragile site and interstitial telomere repeat sequences at the fusion point of a de novo (Y;13) translocation. Hum. Genet. **98**:323-327.

13. Conrad, M.N., A.M. Dominguez, and M.E. Dresser. 1997. Ndj1p, a meiotic telomere protein required for normal chromosome synapsis and segregation in yeast. Science **276**:1252-1255.

14. Cooke, H.J. 1995. Non-programmed and engineered chromosome breakage, p. 219-242. *In* E.H. Blackborn and C.W. Greider (eds.), Telomeres. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

15. Counter, C.M., H.W. Hirte, S. Bacchetti, and C.B. Harley. 1994. Telomerase activity in human ovarian carcinoma. Proc. Natl. Acad. Sci USA **91**:2900-2904.

16. Cuthbert, G., S. McCullough, R. Finney, G. Breese, and N. Brown. 1999. Jumping translocation at 11q23 with MLL gene rearrangement and interstitial telomere sequences. Genes Chromosomes Cancer **24**:295-298.

17. Day, J.P., C.L. Limoli, and W.F. Morgan. 1998. Recombination involving interstitial telomere repeat-like sequences promotes chromosomal instability in Chinese hamster cells. Carcinogenesis **19**:259-265.

18. de Boer, J.G., and B.H. Glickman. 1991. Mutational analysis of the structure and function of the adenine phosphoribosyltransferase enzyme of Chinese hamster. J. Mol. Biol. **221**:163-174.

19. de la Sena, C., B.P. Chowdhary, and I. Gustavsson. 1995. Localization of the telomeric (TTAGGG)n sequences in chromsomes of some domestic animals by fluorescence in situ hybridization. Hereditas **123**:269-274.

20. Desmaze, C., C. Alberti, L. Martins, G. Pottier, C.N. Sprung, J.P. Murnane, and L. Sabatier. 1999. The influence of interstitial telomeric sequences on chromosome instability in human cells. Cytogenet. Cell Genet. **86**:288-295.

21. Dubey, D.D., L.R. Davis, S.A. Greenfeder, L.Y. Ong, J.G. Zhu, J.R. Broach, C.S. Newlon, and J.A. Huberman. 1991. Evidence suggesting that the ARS elements associated with silencers of the yeast mating-type locus HML do not function as chromosomal DNA replication origins. Mol. Cell. Biol. **11**: 5346-5355.

22. Fan, Q.Q., F. Xu, M.A. White, and T.D. Petes. 1997. Competition between adjacent meiotic recombination hotspots in the yeast *Sacchaomyces cerevisiae*. Genetics. **145**:661-670.

23. Farr, C., J. Fantes, P. Goodfellow, and H. Cooke. 1991. Functional reintroduction of human telomeres into mamalian cells. Proc. Nat. Acad. Sci. USA **88**:7006-7010.

24. Farr, C.J., M. Stevanovic, E.J. Thomson, P.N. Goodfellow, and H.J. Cooke. 1992. Telomere-associated chromosome fragmentation: applications in genome manipulation and anlaysis. Nat. Genet. **2**:275-282.

25. Farr, C.J., R.A. Bayne, D. Kipling, W. Mills, R. Critcher, and H.J. Cooke. 1995. Generation of a human X-derived minichromosome using telomere-associated chromosome fragmentation. EMBO J. **14**:5444-5454.

26. Ferguson, B.M. and W.L. Fangman. 1992. A position effect on the time of replication origin activation in yeast. Cell **68**: 333-339.

27. Goldman, A.S., and M. Lichten. 1996. The efficiency of meiotic recombination between dispersed sequences in *Saccharomyces cerevisiae* depends upon their chromosomal location. Genetics **144**:43-55.

28. Gottschling, D.E., O.M. Aparicio, B.L. Billington, and V.A. Zakian. 1990. Position effect at *S. cerevisiae* telomeres: Reversible repression of Pol II transcription. Cell **63**:751-762.

29. Hanish, J.P., J.L. Yanowitz, and T. de Lange. 1994. Stringent requirements for the formation of human telomeres. Proc. Natl. Acad. Sci. USA **91**:8861-8865.

30. Harley, C.B., N.W. Kim, K.R. Prowse, S.L. Weinrich, K.S. Hirsch, M.D. West, S. Bacchetti, H.W. Hirte, C.W. Greider, W.E. Wright, and J.W. Shay. 1994. Telomerase, cell immortality and cancer. *In* Cold Spring Harbor Symposia in Quantitative Biology, Vol. LVIX, pp 307-315. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

31. Hastie, N.D., and R.C. Allshire. 1989. Human telomeres: fusion and interstitial sites. Trends Genet. **5**:326-331.

32. Heller, R., K.E. Brown, C. Burgtorf, and W.R.A. Brown. 1996. Mini-chromosomes derived from the human Y chromosome by telomere directed chromosome breakage. Proc. Natl. Acad. Sci USA **93**:7125-7130.

33. Horn, D., and G.A. Cross. 1995. A developmentally regulated position effect at a terlomeric locus in *Trypanosoma brucei*. Cell **83**:555-561.

34. Ijdo, J.W., A. Baldini, D.C. Ward, S.T. Reeders, and R.A. Wells. 1991. Origin of human chromosome 2: an ancestral telomere-telomere fusion. Proc. Natl. Acad. Sci USA **88**:9051-9055.

35. Itzhaki, J.E., M.A. Barnett, A.B. MacCarthy, V.J. Buckle, W.R. Brown, and A.C. Porter. 1992. Targeted breakage of a human chromosome mediated by cloned human telomeric DNA. Nat. Genet. **2**:283-287.

36. Jewett, T., D. Marnane, W. Stewart, R. Hayworth-Hodge, L. Finklea, K. Klinepeter, P.N. Rao, and M.J. Pettenati. 1998. Jumping translocation with partial duplications and triplications of chromosomes 7 and 15. Clin. Genet. **53**:415-420.

37. Katinka, M.D., and F.M. Bourgain. 1992. Interstitial telomeres are hotspots for illegitimate recombination with DNA molecules injected into the macronucleus of *Paramecium primaurelia*. EMBO J. **11**:725-732.

38. Kim, N.W., M.A. Piatyszek, K.R. Prowse, C.B. Harley, M.D. West, P.L.C. Ho, G.M. Coviello, W.E. Wright, S.L. Weinrich, and J.W. Shay. 1994. Specific association of human telomerase activity with immortal cells and cancer. Science **266**:2011-2015.

39. Kipling, D., H.E. Wilson, E.J. Thomson, M. Lee, J. Perry, S. Palmer, A. Ashworth, and H.J. Cooke. 1996. Structural variation of the pseudoautosomal region between and within inbred mouse strains. Proc. Natl. Acad. Sci. USA **93**:171-175.

40. Kirk, K.E., B.P. Harmon, I.K. Reichardt, J.W. Sedat, and E.H. Blackburn. 1997. Block in anaphase chromosome separation caused by a telomerase template mutation. Science **275**:1478-1481.

41. Kuroiwa, Y., T. Shinohara, T. Notsu, K. Tomizuka, H. Yoshida, S. Takeda, M. Oshimura, and I. Ishida. 1998. Efficient modification of a human chromosome by telomere-directed truncation in high homologous recombination-proficient chicken DT40 cells. Nucl. Acids Res. **26**:3447-3448.

42. Lea, D.E., and C.A. Coulson. 1949. The distribution of the numbers of mutants in bacterial populations. J. Genet. **49**:264-286.

43. Levis, R., T. Hazelrigg, and G.M. Rubin. 1995. Effects of genomic position on the expression of transduced copies of the *white* gene of *Drosophila*. Science **229**:558-561.

44. Luderus, M.E., B. van Steensel, L. Chong, O.C. Sibon, F.F. Cremers, and T. de Lange. 1996. Structure, subnuclear distribution, and nuclear matrix association of the mammalian telomeric complex. J. Cell Biol. **135**:867-881.

45. Luria, S.E., and M. Delbruck. 1943. Mutations of bacteria from virus sensitivity to virus resistance. Genetics **28**:491-511.

46. Merrihew, R.V., R.G. Sargent, and J.H. Wilson. 1995. Efficient modification of the APRT gene by FLP/FRT site-specific targeting. Som. Cell & Mol. Genet. **21**:299-307.

47. Merrihew, R.V., K. Marburger, S.L. Pennington, D.B. Roth, and J.H. Wilson. 1996. High-frequency illegitimate integration of transfected DNA at preintegrated target sites in a mammalian genome. Mol. Cell. Biol. **16**:10-18.

48. Miklos. G.L.G. and R.N. Nankivell. 1976. Telomeric satellite DNA functions in regulating recombination. Chromosoma **56**:143-167.

49. Mills, W., R. Critcher, C. Lee, and C.J. Farr. 1999. Generation of an approximately 2.4 Mb human X cetromere-based minichromosome by targeted telomere-associated chromosome fragmentation in DT40. Hum. Mol. Genet. **8**:751-761.

50. Musio, A., G. Rainaldi, and I. Sbrana. 1996. Spontaneous and aphidicolin-sensitive fragile site 3cen co-localizes with the (TTAGGG)n telomeric sequence in Chinese hamster cells. Cytogenet. Cell Genet. **75**:159-163.

51. Nairn, R.S., R.M. Humphrey, and G.M. Adair. 1988. Transformation of UV-hypersensitive Chinese hamster ovary cell mutants with UV-irradiated plasmids. Int. J. Radiat. Biol. **53**:249-260.

52. Nimmo, E.R., G. Cranston, and R.C. Allshire. 1994. Telomere-associated chromosome breakage in fission yeast results in variegated expression of adjacent genes. EMBO J. **13**:3801-3811.

53. Park, V.M., T.M. Gustashaw, and T.M. Wathen. 1992. The presence of interstitial telomeric sequences in constitutional chromosome abnormalities. Am. J. Hum. Genet. **50**:914-923.

54. Prado, F., and A. Agulera. 1995. Role of reciprocal exchange, one-ended invasion crossover and single-strand annealing on inverted and direct repeat recombination in

yeast: Different requirements for the *RAD1*, *RAD10*, and *RAD52* genes. Genetics **139**:109-123.

55. Regad, F., M. Lebas, and B. Lescure. 1994. Interstitial telomeric repeats within the *Arabidopsis thaliana* genome. J. Mol. Biol. **239**:163-169.

56. Reynolds, A.E., R.M. McCarroll, C.S. Newlon, and W.L. Fangman. 1989. Time of replication of ARS elements along yeast chromosome III. Mol. Cell. Biol. **9**:4488-4494.

57. Rossi, E., G. Floridia, M. Casali, C. Danesino, G. Chiumello, F. Bernardi, I. Magnani, L. Papi, M. Mura, and O. Zuffardi. 1993. Types, stability, and phenotypic consequences of chromosome rearrangement leading to interstitial telomeric sequences. J. Med. Genet. **30**:926-931.

58. Roth, D.B., and J.H. Wilson. 1988. Illegitimate recombination in mammalian cells, p. 621-654. *In* R. Kucherlapati and G.R. Smith (eds.), Genetic recombination. American Society for Microbiology, Washington, D.C.

59. Rouet, P., F. Smih, and M. Jasin. 1994. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. Proc. Natl. Acad. Sci USA **91**:6064-6068.

60. Rudenko, G., P.A. Blundell, A. Dirks-Mulder, R. Kieft, and P. Borst. 1995. A ribosomal DNA promoter replacing the promoter of a telomeric VSG gene expression site can be efficiently switched on and off in *T. brucei*. Cell **83**:547-553.

61. Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989 Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

62. Sargent, R.G., R.V. Merrihew, R. Nairn, G. Adair, M. Meuth, and J.H. Wilson. 1996. The influence of a (GT)29 microsatellite sequence on homologous recombination in the hamster APRT gene. Nucl. Acids Res. **24**:746-753.

63. Sargent, R.G., M. Brenneman, and J.H. Wilson. 1997. Repair of site-specific double-strand breaks on a mammalian chromosome by homologous and illegitimate recombination. Mol. and Cell Biol. **17**:267-277.

64. Sargent, R.G., R.L. Rolig, A.E. Kilburn, G.M. Adair, J.H. Wilson, and R.S. Nairn. 1997. Recombination-dependent deletion formation in mammalian cells deficient in the nucleotide excision repair gene ERCC1. Proc. Natl. Acad. Sci. USA **94**:13122-13127.

65. Slijepcevic, P., Y. Xiao, A.T. Natarajan, and P.E. Bryant. 1997. Instability of CHO chromosomes carrying interstitial telomere sequences originating from Chinese hamster chromosome 10. Cytogenet. Cell Genet. **76**:58-60.

66. Slijepcevic, P. 1998. Telomeres and mechanisms of Robertsonian fusion. Chromosoma **107**:136-140.

67. Stavenhagen, J.B., and V.A. Zakian. 1998. Yeast telomeres exert a position effect on recombination between internal tracts of yeast telomeric DNA. Genes & Dev. **12**:3044-3058.

68. Stevenson, J.B. and D.E. Gottschling. 1999. Telomeric chromatin modulates replication timing near chromosome ends. Genes Dev. **13**:146-151.

69. Tommerup, H., A. Dousmansis, and T. de Lange. 1994. Unusual chromatin in human telomeres. Mol. Cell. Biol. **14**:5777-5785.

70. van Steensel, B., A. Smorgorzewska, and T. de Lange. 1998. TRF2 protects human telomeres from end-to-end fusions. Cell **92**:401-413.

71. Vermeesch, J.R., P. Petit, F. Speleman, K. Deuriendt, J.P. Fryns, and P. Marynen. 1997. Interstitial telomeric sequences at the junction site of a jumping translocation. Hum. Genet. **99**:735-737.

72. Wang, P., R.H. Zhou, Y. Zou, C.K. Jackson-Cook, and L.F. Povirk. 1997. Highly conservative reciprocal translocations formed by apparent joining of exchanged DNA double-strand break ends. Proc. Natl. Acad. Sci. USA **94**:12018-12023.

73. Xu, F., and T.D. Petes. 1996. Fine-structure mapping of meiosis-specific double-strand DNA breaks at a recombination hotspot associated with an insertion of telomeric sequences upstream of the HIS4 locus in yeast. Genetics **143**:1115-1125.

Table 1. Relative transfection efficiency of *APRT* genes with telomere sequence.

Plasmid Construct	**Transfection Method**	***APRT*$^+$ Frequency**	***GPT*$^+$ Frequency**	***APRT*$^+$/*GPT*$^+$ Ratio**
pGS36 (No Insert)	Electroporation	3.3×10^{-5}	4.0×10^{-5}	0.8
pAK301 (CCCTAA)	Electroporation	1.0×10^{-4}	1.0×10^{-4}	1.0
pAK501 (TTAGGG)	Calcium Phosphate	4.0×10^{-3}	5.4×10^{-3}	0.7

Table 2. Rates of *TK⁻APRT⁻* and *APRT⁻* colony formation and analysis of independent products

A *TK⁻APRT⁻*

Cell Line	Insert	Rate (x 10^{-7})[b]	Products[c]				
			Crossover	Convers.	Rearran.	Mutation	Total
GSB, C, E[a]	<200	1.2 ± 0.2	72	0	0	0	72
AK209	HPRT	1.2 ± 0.4	12	0	0	0	12
AK92, AK213	CCCTAA	1.9 ± 0.5	11	0	2	0	13
AK775	TTAGGG(I-SceI)	1.3	9	0	1	0	10

B *APRT⁻*

Cell Line	Insert	Rate (x 10^{-7})[b]	Products[c]				
			Crossover	Convers.	Rearran.	Mutation	Total
GSB, C, E[a]	<200	15 ± 1	17	88	0	6	111
AK209	HPRT	14 ± 4	0	11	0	0	11
AK92, AK213	CCCTAA	14 ± 3	2	12	3	7	24
AK775	TTAGGG(I-SceI)	16	2	6	4	0	12

[a]Rates and product analysis for the GSB, GSC, and GSE cell lines are from previous reports (62, 64). GSB cell lines contained a 114-bp insert carrying recognition sequences for the HO endonuclease and for the FLP recombinase; GSC cell lines carried no insert; GSE cell lines carried a 200-bp insert that contained a $(GT)_{29}$ repeat.

[b]Rates were determined by fluctuation analyses. For AK209 the rate is an average of two fluctuation analyses. One analysis for AK92 and two for AK213 were averaged. The rate for AK775 is from a single fluctuation analysis. The standard error of the mean is indicated.

[c]The molecular structure of the APRT locus in individual colonies was determined by a combination of Southern and PCR analysis.

Table 3. *TK⁻APRT⁻* and *APRT⁻* colony formation after treatment with I-SceI endonuclease.

Cell Line	Insert	*TK⁻APRT⁻* (x 10^{-4})[a]	*APRT⁻* (x 10^{-4})[a]
GSAA5	I-SceI	1.3 ± 0.3	18 ± 7
AK846	HP(I-SceI)RT	0.8 ± 0.2	9 ± 3
AK696	(I-SceI)CCCTAA	5.7 ± 0.6	10 ± 2
AK775	TTAGGG(I-SceI)	4.3 ± 0.2	6 ± 2

[a]The values for GSAA5 are the average of six experiments; the values for the other cell lines are the average of three experiments. The standard error of the mean is indicated. Stimulation above spontaneous frequencies of *TK⁻APRT⁻* and *APRT⁻* cells in untreated populations averaged about 100-fold for each selection.

Figure Legends:

Figure 1. Molecular structures of the substrates at the *APRT* locus in diploid and tetraploid cell lines and of the products isolated in various selections. (A) Inserted sequences are shown above their common site in the second intron of the downstream, functional *APRT* gene (the five exons of *APRT* are shown as boxes). Vertical arrows indicate locations of the I-SceI recognition sites. Inverted triangles indicate the positions of the FRT recognition sequences. The upstream copy of *APRT* is nonfunctional by virtue of a truncated fifth exon and a mutation in exon 2 (filled box). The upstream and downstream copies share 6.8 kb of homology; 4.5 kb upstream of the APRT gene (thick line) and 2.3 kb of homology within the gene itself. Cleavage sites for the restriction enzymes BamHI (B), HindIII (H), and BclI (Bc) that were used in Southern analyses are indicated. An additional BamHI site in AK775 and AK858 is located in the polylinker adjacent to the TA-end of the telomere sequence (not shown). The hybridization probe corresponds to the downstream BamHI fragment that encompasses the *APRT* gene, but included no inserted sequences. (B) Products were distinguished based on Southern patterns after BamHI and HindIII cleavage and PCR analysis (10-12). Conversions have a structure like the parental tandem duplication, except that some lose the insert as part of the conversion process (status of the insert is indicated by +/–). Conversions were shown to contain the EcoRV mutation in exon 2 (filled box) by PCR amplification across the exon (see Figure 4A) followed by incubation with EcoRV. Crossovers have a single copy of the *APRT* gene whose size depends on whether the insert was retained or lost. Rearrangements yield a Southern pattern that does not correspond to conversions or crossovers (see Figure 3); they were subjected to further Southern and PCR analysis (see

Figure 4A). Mutations were identical to conversions by Southern analysis, but were shown not to contain the EcoRV mutation by PCR analysis. They are assumed to carry point mutations or small deletions elsewhere in the *APRT* gene; however, they have not been further characterized.

Figure 2. Northern analysis of *APRT* expression in various cell lines. Lane 1, transcript levels in cell line RMP41 (20), which has a single copy of the *APRT* gene. Lane 2, transcript levels in cell line DELI26 (20), which has a tandem duplication structure but the downstream copy of *APRT* is lacking the promoter and 5′ half of the gene. Absence of a transcript in this strain indicates that no stable transcript is made from the truncated, upstream copy of *APRT*, which is identical to that in the other strains tested. Lane 3, transcript levels in cell line GSAA5, which carries I-SceI and FRT recognition sites. Lane 4, transcript levels in cell line AK209, which carries an *HPRT* insert. Lane 5, transcript levels in cell line AK213, which carries a CCCTAA insert. Lane 6, transcript levels in cell line AK775, which carries a TTAGGG insert. Hybridization to GAPDH serves as a loading control.

Figure 3. Southern analysis of *TK⁻APRT⁻* and *APRT⁻* colonies from AK775. DNAs from individual colonies were digested with BamHI and HindIII, and the fragments were resolved by electrophoresis and visualized by Southern blotting using a ^{32}P-labelled BamHI fragment of the wild type *APRT* gene as a probe. Crossovers that have lost the telomere sequence have a single band at 4.0 kb; crossovers that have retained the telomere sequence have bands at 3.5 kb and 1.3 kb. There is a BamHI site (not shown in

Figure 1A) at the TA end of the telomere sequence, so that the telomere sequence is in the 3.5 kb band. Conversions have the 7.0 kb band from the upstream copy of *APRT*. Conversions that have lost the telomere sequence have an additional band at 4.0 kb; conversions that retain the telomere sequence have two additional bands at 3.5 kb and 1.3 kb. Rearrangements have patterns that do not match these expectations; their identities are indicated at the top of the autoradiograph. Numbers at the sides indicate the length of fragments in kb. RMP41 carries a single copy of the *APRT* gene; AK92 carries a tandem duplication with telomere sequence in the CCCTAA orientation (it does not carry a BamHI site adjacent to the telomere sequence).

Figure 4. Molecular structures of rearrangements and sequences of some rearrangement junctions. (A) Locations of restriction-enzyme and PCR-primer sites are shown in the tandem duplication structure at the top. Parent cell lines from which the rearrangements arose are shown in brackets at left. Sizes of deletions and insertions are estimates based on Southern and PCR analysis. Crossovers and conversions were confirmed by lack of cleavage of the PCR products across the second exon, indicating the presence of the EcoRV mutation. AK613 is designated a translocation because it gives two bands upon cleavage by BamHI, HindIII, or BclI; however, it could be an insertion of DNA that includes all three recognition sites. (B) Nucleotide sequences around the insertion point for the parental cell lines, AK213 and AK775, are shown along with the sequences of the rearrangement junctions. TTGGGG sequences (bold) are interspersed with TTAGGG sequences (normal) and one TTAGCG sequence (italics) at the nonseeding (TA-rich) end of the telomere sequence. Dashes indicated the large number of predominantly

TTAGGG sequences toward the seeding (G-rich) end of the telomere sequence. The site of insertion in *APRT*, nucleotide 1310, is indicated, as are the last nucleotides of *APRT* sequences that flank the insertion in AK586.

Figure 1

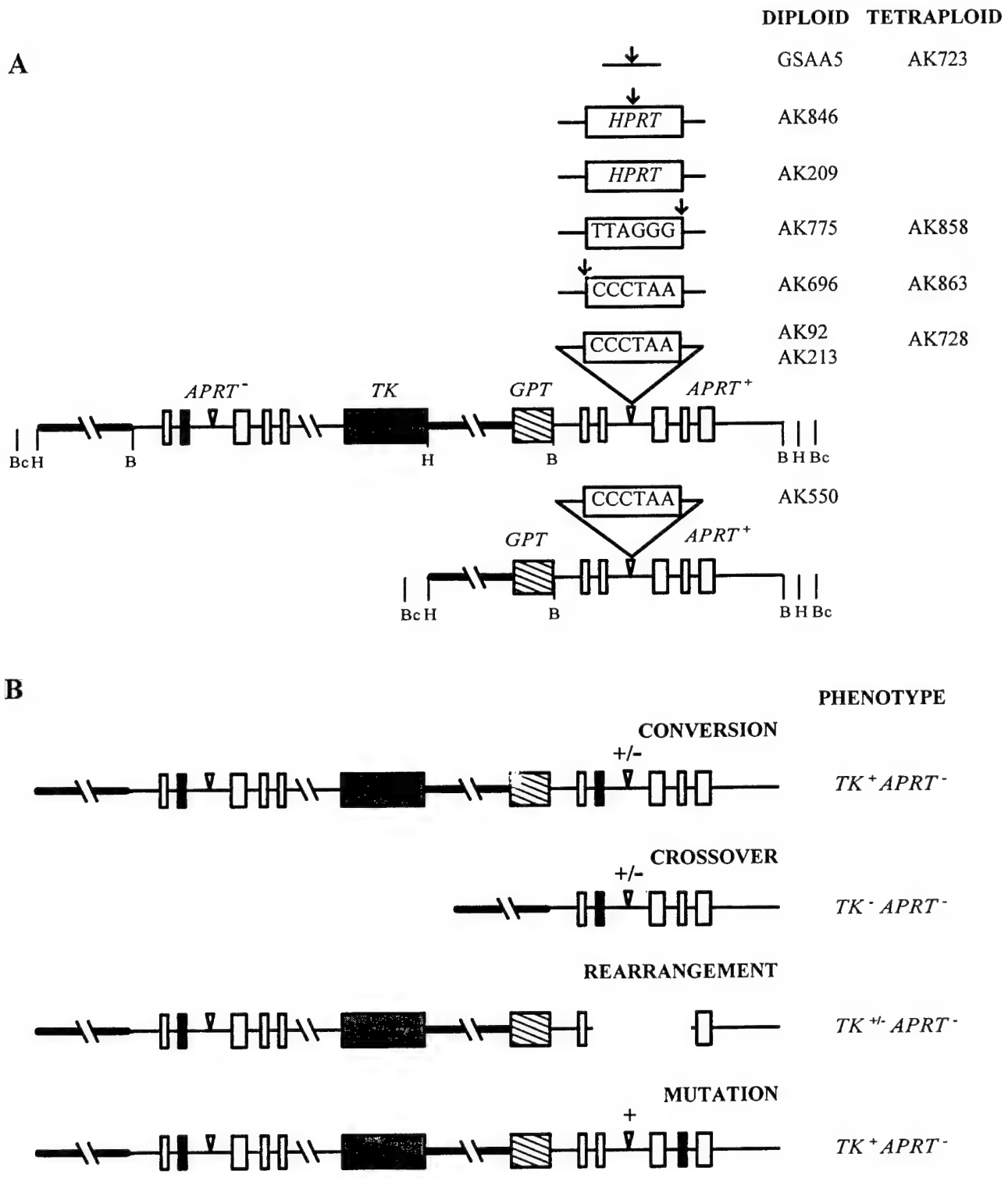
DIPLOID
TETRAPLOID
A
GSAA5
AK723
HPRT
AK846
HPRT
AK209
TTAGGG
AK775
AK858
CCCTAA
AK696
AK863
CCCTAA
AK92
AK213
AK728
APRT-
TK
GPT
APRT+
BcH
B
H
B
B H Bc
CCCTAA
AK550
GPT
APRT+
BcH
B
B H Bc
B
PHENOTYPE
CONVERSION
+/-
TK+APRT-
CROSSOVER
+/-
TK-APRT-
REARRANGEMENT
TK+/-APRT-
MUTATION
+
TK+APRT-

Figure 2

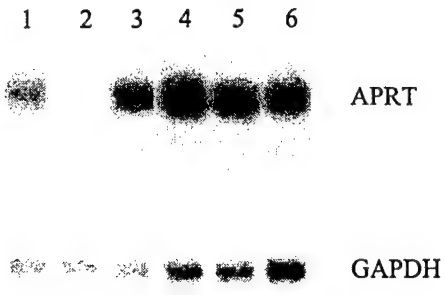

Figure 3

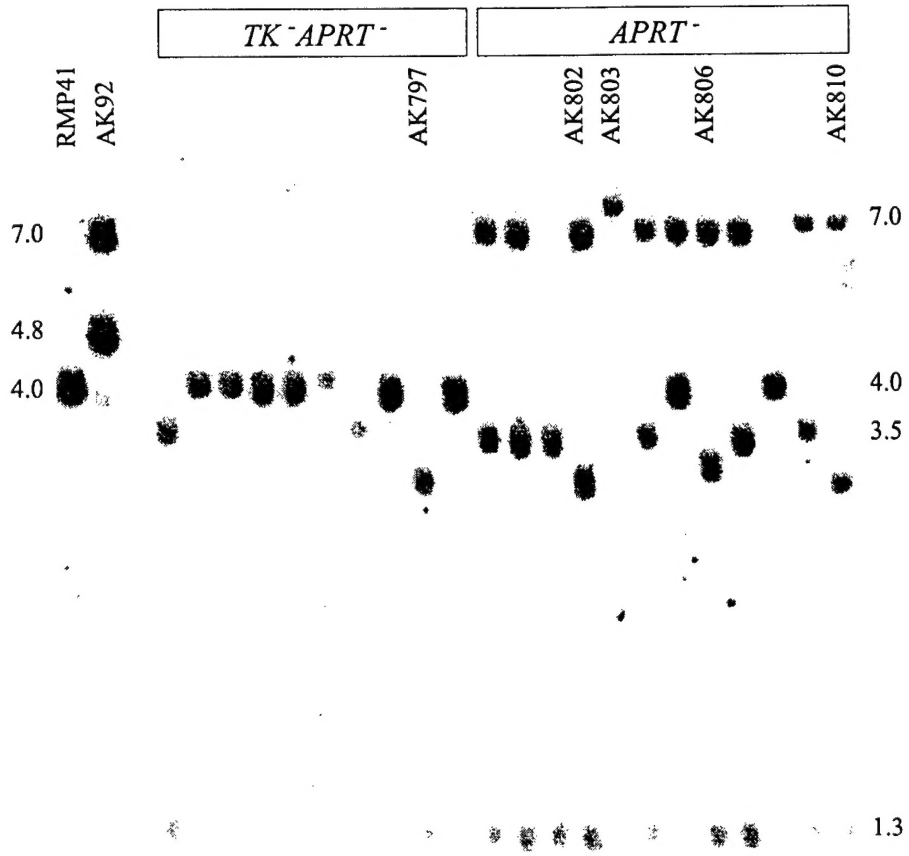
TK⁻APRT⁻
APRT⁻
RMP41
AK92
AK797
AK802
AK803
AK806
AK810
7.0
4.8
4.0
7.0
4.0
3.5
1.3

Figure 4

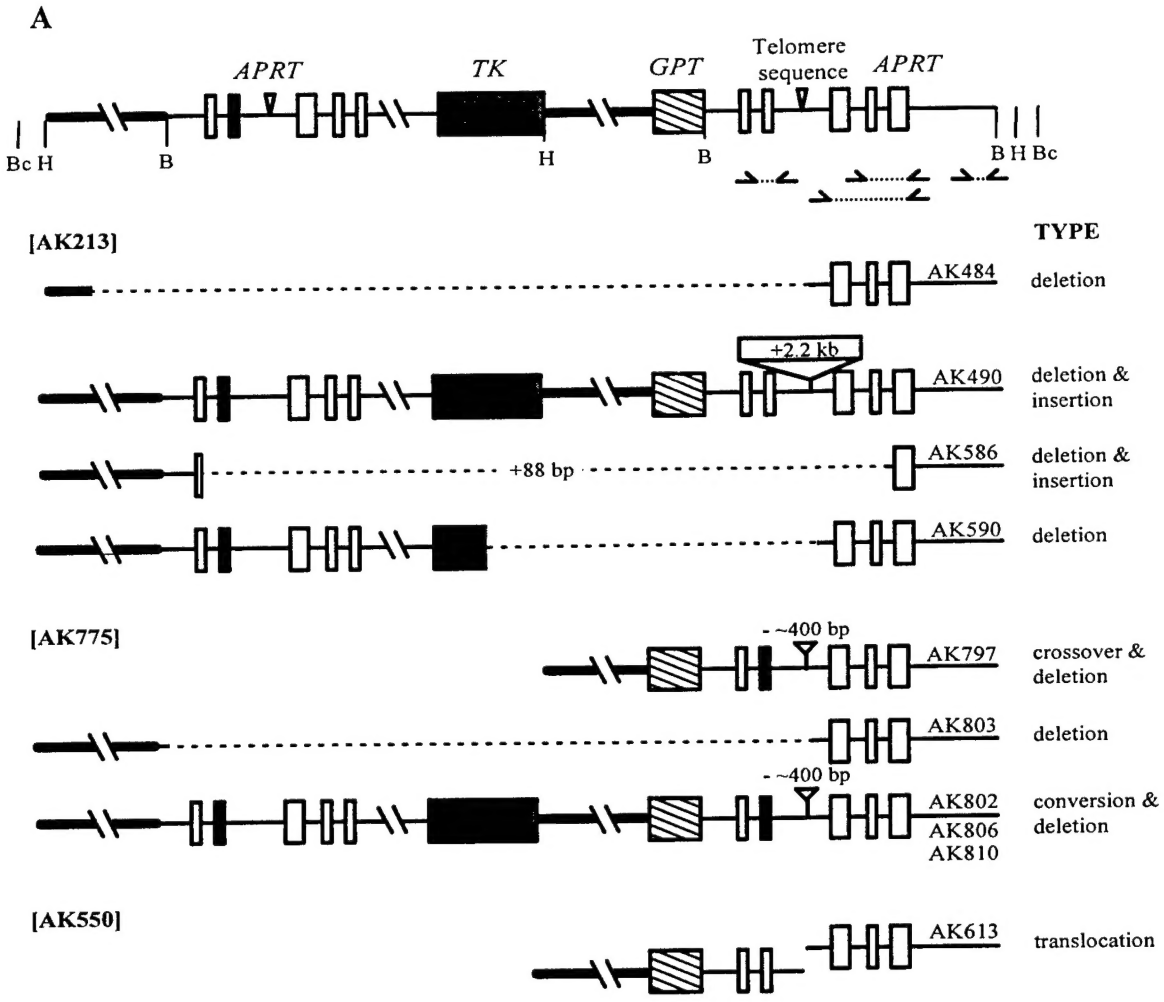
A
APRT
TK
GPT
Telomere sequence
APRT
Bc H
B
H
B
B H Bc
[AK213]
TYPE
AK484
deletion
+2.2 kb
AK490
deletion & insertion
+88 bp
AK586
deletion & insertion
AK590
deletion
[AK775]
~400 bp
AK797
crossover & deletion
AK803
deletion
~400 bp
AK802
AK806
AK810
conversion & deletion
[AK550]
AK613
translocation

Figure 4B

[AK213]

APRT(1310) poly GAATTCCGAA CCCTAA ------ CCCTAA **CCCCAA CCCCAA CCCCAA CCCCAA CCCCAA CCCCAA CCCCAA CCCCAA** CCCTAA CCCTAA CCCTAA CCCTAA CCCTAA CCCTAA **CCCCAA** CCCTAA *CGCTAA* CCCTAA **CCCCAA** CCCTAA **CCCCAA CCCCAA CCCCAA CCCCAA CCCCAA** CCGGAATTC poly FRT poly (1311)*APRT*

AK586 411

GGAATCTGAGTTG TTAGGG TTAGGG TTAGGG TTAGGG TTAGGG **TTGGGG TTGGGG** TTAGGG TTAGGG TTAGGG **TTGGGG TTGGGG TTGGGG** *TTCAGGG* TTA GGGGTGCTGA

2104

[AK775]

APRT(1310) poly GAATTCCGG **TTGGGG TTGGGG TTGGGG TTGGGG TTGGGG** TTAGGG **TTGGGG** TTAGGG *TTAGCG* TTAGGG **TTGGGG** TTAGGG TTAGGG TTAGGG TTAGGG TTAGGG TTAGGG **TTGGGG TTGGGG TTGGGG TTGGGG TTGGGG TTGGGG TTGGGG TTGGGG** TTAGGG ------ TTAGGG TTCGGAATTC poly I-SceI poly FRT poly (1311)*APRT*

AK797, AK802, AK806, AK810

APRT(1310) poly GAATTCCGG **TTGGGG TTGGGG TTGGGG TTGGGG TTGGGG** TTAGGG **TTGGGG** TTAGGG *TTAGCG* TTAGGG **TTGGGG** TTAGGG TTAGGG TTAGGG TTAGGG TTAGGG TTAGGG ------ TTAGGG TTCGGAATTC poly I-SceI poly FRT poly (1311)*APRT*

Effects of Telomere Sequence on the Stability and Expression of a Defined Gene in Mammalian Cells

April E. Kilburn*, Martin J. Shea†, R. Geoffrey Sargent*, Zoltan Sandor*, and John H. Wilson*†
*The Verna and Marrs McLean Department of Biochemistry and Molecular Biology
†The Department of Molecular and Human Genetics
Baylor College of Medicine, Houston TX 77030

Telomere repeat sequences help stabilize the ends of chromosomes; however, at interstitial sites they are correlated with genetic instability and the phenomenon of telomere-associated chromosome fragmentation (TACF). To determine the effects of interstitial telomere sequence on genome stability, we deposited 800 bp of the vertebrate telomere repeat, TTAGGG, in two orientations in the second intron of the adenosine phosphoribosyltransferase (*APRT*) gene in Chinese hamster ovary cells. The telomere sequence did not interfere with expression of the *APRT* gene. By selection for loss of APRT function, we measured the ability of the telomere sequence to stimulate homologous and illegitimate recombination and to cause chromosomal truncation. Although telomere sequence did not alter the rate or distribution of homologous recombination events, it dramatically increased local illegitimate recombination events some 30-fold to a frequency between 10^{-6} and 10^{-7}. Analysis of individual rearrangements confirmed the involvement of the telomere sequence. Using tetraploid cells to provide a second copy of the partially hemizygous APRT chromosome, we showed that interstitial telomere sequence did not induce chromosome truncation by the seeding of new telomere at a frequency greater than 0.1%. Such a low frequency of truncation is much below the 20-70% frequencies reported for chromosomal truncation in TACF experiments, suggesting that TACF does not occur by integration of the plasmid followed by breakage at the telomere sequence. It is also below the detection limits for the cytogenetic assays that have been used to assign instability to naturally occurring, but uncharacterized, interstitial telomere sequences. Instability at natural interstitial sites may relate to their much larger size, usually several kilobases, or to details of their structure. The relative stability of short telomere sequence offers the possibility for constructing defined tester chromosomes to measure the effects of destabilizing treatments, stability in different genetic backgrounds, and the influence of oncogenic transformations.